Journal of Cleaner Production 173 (2018) 143–150

Contents lists available at ScienceDirect

Journal of Cleaner Production

journal homepage: www.elsevier.com/locate/jclepro

A novel method for carbon dioxide emission forecasting based on improved Gaussian processes regression

Debin Fang [a], Xiaoling Zhang [b], Qian Yu [c, *], Trenton Chen Jin [d], Luan Tian [d]

[a] *Economics and Management School, Wuhan University, Wuhan, 430072, China*
[b] *Department of Public Policy, City University of Hong Kong, Hong Kong, China*
[c] *School of Economics, Wuhan University of Technology, Wuhan, 430070, China*
[d] *Economics Department, The George Washington University, Washington DC, 20052, United States*

ARTICLE INFO

Article history:
Received 6 July 2016
Received in revised form
14 December 2016
Accepted 19 May 2017
Available online 28 May 2017

Handling Editor: Yutao Wang

Keywords:
CO2 emission
Gaussian process regression
Particles swarm optimization
Forecasting

ABSTRACT

The fact that global warming will bring impact on immigration, agriculture and also generate human conflicts is becoming a focus in climate change topic and the forecasting of carbon dioxide emission has been attracting much attention. In this paper, we proposed an improved Gaussian processes regression method for carbon dioxide emission forecasting based on a modified PSO algorithm which can efficiently optimize the hyper parameters of covariance function in the Gaussian processes regression. Also we tested our improved PSO-GPR method with the total carbon dioxide emissions data of U.S., China and Japan in 1980–2012, and compared the prediction precision of our method with original GPR and BP Neural Networks by the data of U.S., China and Japan. The performance of our improved Gaussian processes regression method enhanced the prediction accuracy of original GPR method and is superior to other traditional forecasting method like BP Neural Networks. Furthermore, we applied PSO-GPR method in generating a prediction total carbon dioxide emissions for 2013 to 2020 and found out that China's total carbon dioxide emission will still increasing but finally at a decreasing rate and U.S. and Japan will have a good control on their amount of carbon emission in the near future. Finally, policy implications about carbon dioxide emission reduction were proposed.

1. Introduction

As we marched into the new millennium, climate change has become unequivocal and exerted various unprecedented impacts on all walks of life. Greenhouse Gases (GHGs) is the dominant cause of global warming which influences diversified human economic activities (Nordhaus and William, 2006), such as agriculture yields (Robert et al., 1994; Deschenes et al., 2007), industrial productivity (Dell et al., 2013), populations (Andrew et al., 2010) and immigration (Feng et al., 2010).

Drastic climate variations like rising temperature, diminishing ice and increased sea level, will inevitably give rise to destruction to ecosystem, biodiversity and human economic activities both in the short run and in the long run. The deteriorating environment quality and vanishing species harshly make us realize the necessity to take actions against any further degradation. The irreversible damages require global joint forces to deal with these urgent situations and maintain sustainable development. Therefore, one crucial question with regard to policy is whether we should exert efforts in GHGs emission reduction? However, one main issue in the process is to accurately predict GHGs emission level. Despite the fact that GHGs projection is a challenging mission, it is an imperative and worthwhile path to pursue.

Future scientific and economic uncertainties have also been neglected when estimating climate change in the previous literature. Stern and Nicholas (2008) firstly conducted a cost-benefit analysis to investigate the necessity of GHGs emission mitigation under an expected utility framework. The paper opens up the floodgate of a series of researches that make strong arguments about GHGs emission reduction. Both developed and developing countries have witnessed falling agriculture yields, increased desertification over these decades. It seems that countries around the world reach a consensus as to fight against climate change is benefit for GHGs emission reduction.

* Corresponding author.
E-mail addresses: dbfang@whu.edu.cn (D. Fang), xiaoling.zhang@cityu.edu.hk (X. Zhang), yuqian@whut.edu.cn (Q. Yu), trentonjin@gwu.edu (T.C. Jin), luantian0621@gwu.edu (L. Tian).

http://dx.doi.org/10.1016/j.jclepro.2017.05.102
0959-6526/

As to 2012, the actual global emissions reached 34.5 billion tons, equivalent to 1.0104 times the amount of 2011(Trends in Global CO2 Emissions: Report 2013, 2013). The remarkable upward trends in carbon emissions mainly attribute to 3 regions which accounted for 55% global carbon dioxide emissions. Among the top 3 emitting regions, China, United States (Mattinen et al., 2014) and the European Union (Dragomir, 2012) take up 29%, 16%, 11% respectively. The research results of Liu et al. (2010) and Geng et al. (2013), which are about China's regional carbon emissions change over 1997–2007 and the driving factors of energy-related carbon dioxide emissions in Chinese provinces (Miao et al., 2016), have shown that the surge in carbon emissions is mainly related to energy consumption in most provinces and regions of China as developing countries (Sun and Liu, 2016). Also Kyle and Juliet (2014) has studied the economic growth and climate change in high-income countries by cross-national analysis. These studies indicated that the continuing trend will impose threat to sustainable development of developing and impede the process to achieve emission reduction targets (Feng et al., 2010).

For forecasting carbon dioxide emission, a number of energy consumption-based greenhouse gas emission models have been developed (Morita et al., 1994). Mikiko et al. (2000) used The Asian-Pacific Integrated Model (AIM), based on differences socioeconomic scenarios, to predict greenhouse gas emissions in Japan, and it found that it is possible to mitigate carbon dioxide emissions without scaling back productive activities or standards of living. In recently, Genetic Algorithm (GA) has been used to forecast carbon dioxide emission based on the global oil (Reijnders and Huijbregts, 2008), natural gas, coal, and primary energy consumption figures by Kavoosi et al. (2012). Sun (2013) presented a Harmony Search based optimization grey forecasting model to forecast carbon dioxide emissions in China. Abdel (2013) proposed an Artificial Neural Network model (ANN) with four inputs data including global oil, natural gas, coal, and primary energy consumption to handling the time series forecasting of carbon dioxide emissions. However, ANN may lead to complex network structure and substantial learning time in forecasting of carbon dioxide emission. In addition, despite the strong nonlinear modeling capacity, these models mentioned above cannot exactly characterize the feature of small sample size and high-dimension data of carbon dioxide emission prediction. To address these issues, Gaussian process regression (GPR) approach is developed to tackle the small sample and high-dimension data effectively and more accurately for prediction of nonlinear.

GPR's flexibility, nonlinearity and inherent nonparametric structure made it gained popularity across various application fields such as chemistry (Yu et al., 2013a,b), astrophysics (Remya et al., 2015), material (Piero et al., 2015). Martin et al. (2014) used Gaussian process regression for the analysis of electroanalytical experimental data to estimate diffusion coefficients from voltammetric signals. Chen et al. (2013) applied GPR to optimal design of combustion systems that using real-time flame images to predict the outlet content of the flue gas. In addition, GPR is frequently applied to wind speed prediction (Kou et al., 2013). Yu et al. (2013a,b) proposed a Gaussian mixture copula model based localized GPR approach for predicting long-term wind speed. A combination model that integrates the ELM, SVM, LSSVM and GPR models for probabilistic short-term wind speed forecasting was presented by Wang and Hu (2015). And to solve the problem of large wind data sets computing, Arash et al. (2014) presents an approximation algorithm called Bayesian site selection method to reduce the computation time of GPR. According to the characteristics of carbon dioxide missions, it is suitable to adopt GPR to forecast carbon dioxide emissions. However, there are almost no related researches in this field until now.

In this paper, we firstly introduce GPR to predict carbon dioxide emissions, however, widespread availability of data, while furnishing adequate information to train the model, could hamper computationally efficient implementation of GPR (Arash et al., 2014). For the purpose of improving the accuracy of prediction and efficiently of computation, we proposed an improved GPR method based on a modified Particle Swarm Optimization(PSO) algorithm. Optimize the parameters based on the sample learning is a key characteristics in GPR. By integrating the GPR method with a modified PSO algorithm, the hyper parameters of covariance function in the GPR model can be improved optimization efficiency. Consequently, the performance of the original GPR can be improved.

Section 2 introduces the Gaussian progress regression method with a modified PSO algorithm for optimization of hyper parameters. The optimization of hyper parameters can reduce the computation redundancy and enhance estimation performance. Section 3 describes the data resources and estimation models. We seek to estimate carbon dioxide emission levels for U.S., China and Japan. We will test our improved PSO-GPR method with the total carbon dioxide emissions data of U.S., China and Japan in 1980–2012, and compared the prediction precision of our method with original GPR and BP Neural Networks by the data of U.S., China and Japan. Furthermore, we apply PSO-GPR method in generating a prediction total carbon dioxide emissions for 2013 to 2020 and analysis the policy implications about carbon dioxide emission reduction. Conclusions will be given out in Section 4.

2. Methodology

2.1. Preliminary and notation of Gaussian process regression (GPR)

A Gaussian process (GP) is a collection of random variables, any finite number of which has consistent joint Gaussian distributions. A Gaussian process can be completely specified by its mean function and covariance function and can be denoted as $f(x) \sim GP(m(x), C(x, x'))$, where x is the carbon dioxide (CO_2) emission, $m(x) = E[f(x)]$ is the mean function and $C(x, x') = E[(f(x) - m(x))(f(x') - m(x'))]$ is the covariance function (Rasmussen and Williams, 2006).

In general, positive semi-definite kernels are valid covariance functions. A very common covariance function is the squared-exponential covariance function (Tobias et al., 2006), in one dimension has the following form: $C(x_p, x_q) = \theta^2 \exp\left(-\frac{\|x_p - x_q\|^2}{2l^2}\right)$, where θ controls the prior variance, and l is an isotropic length scale parameter that controls the rate of decay of the covariance, i.e. determines how far away x_p must be from x_q for f_p to be unrelated to f_q. For this particular covariance function, we see that the covariance is almost unity between variables whose corresponding inputs are very close, and decreases as their distance in the input space increases.

Suppose we are given a training dataset $D = \{x^{(i)}, y^{(i)} | i = 1, 2, \cdots n\}$ with noise observation, and then the method of Gaussian process regression (GPR) can be employed to predict the target value y_* for the new input value x_* by learning a

function from the dataset, which involves an assumed Gaussian prior of functions.

According to Bayer's rule, we can obtain the posterior distribution for the $(n+1)$ Gaussian process outputs if we compute the distribution of the new point. Given a test input $x^{(n+1)}$, and the set of training points $D = \{x^{(i)}, y^{(i)} | i = 1, 2, \cdots n\}$. By conditioning on the observed targets in the training set, the predictive distribution is Gaussian $P(y^{(n+1)} | D, x^{(n+1)}) \sim N(\mu_{y^{(n+1)}}, \sigma^2_{y^{(n+1)}})$ where $\mu_{y^{(n+1)}} = a^T Q^{-1} y$ is the mean, and $\sigma^2_{y^{(n+1)}} = C(x^{(n+1)}, x^{(n+1)}) - a^T Q^{-1} a$ is variance respectively, given $Q_{pq} = C(x^{(p)}, x^{(q)}) + r^2 \delta_{pq}$, $a_p = C(x^{(n+1)}, x^{(q)})$, $p = 1, \cdots, n$.

As discussed above, covariance function $C(x_p, x_q; \Theta)$ and its hyper parameters Θ played an important role in Gaussian process regression because it controls how much the data are smoothed in estimating the unknown function. These parameters should to be set properly for better learning experience.

For optimize the hyper parameters from the training data, we can maximize the log likelihood $\log P(D|\Theta) = \log P(y^{(1)}, \cdots, y^{(n)} | x^{(1)}, \cdots, x^{(n)}, \Theta) = -\frac{1}{2} \log \det C - \frac{1}{2} y^T C^{-1} y - \frac{n}{2} \log 2\pi$ of the hyper parameters by introduce an improved Particles Swarm Optimization (PSO) algorithm into GPR method.

2.2. Modification of particles swarm optimization(PSO)

PSO is an algorithm proposed by Kennedy and Eberhar (1995), which utilize the swarm intelligence generated by the cooperation and competition between the particles in a swarm and has emerged as a useful tool for global optimization.

In a PSO algorithm, each particle, standing for a candidate solution, flies in a D-dimensional space S according to the historical experiences of its own and its colleagues. And the quality of solution is determined by the fitness function. The position and velocity of the ith particle is represented by $x_i = (x_{i1}, \ldots, x_{id}, \ldots, x_{iD})$ and $v_i = (v_{i1}, \ldots, v_{id}, \ldots, v_{iD})$ respectively. PSO algorithm randomly initializes a swarm of particles and finds the optimized solution via randomly searching. In iteration, each particle keeps track of its coordinates in hyperspace according to two "fitness values" it has achieved so far: The optimum position of individual particle i is stored as $p_i = (p_{i1}, p_{i2}, \cdots, p_{iD})$, denoted as *pbest* and the best position of the swarm can be recorded as $p_g = (p_{g1}, p_{g2}, \cdots, p_{gD})$, which is *gbest.* After getting the *pbest* and *gbest*, the velocity and position of the i th particle can be updated according to following equations:

$$v_{id}(t+1) = w v_{id}(t) + c_1 r_1 [p_{id} - x_{id}(t)] + c_2 r_2 \left[p_{gd} - x_{id}(t)\right] \quad (1)$$

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1) \quad (2)$$

where r_1 and r_2 are two uniform random sequences sampled from $U(0,1)$, c_1 and c_2 are two positive acceleration coefficients, which are commonly set as $c_1 = c_2 = 2$, and w is an inertia weight determining how much of the particle's previous velocity is preserved. The value of w decrease linearly with the iteration as follows: $w = w_{\max} - iter*\frac{w_{\max} - w_{\min}}{iter_{\max}}$, $w_{max} = 0.9$, $w_{min} = 0.4$.

In order to maintain the diversity and avoid the problem of prematurity and easily trapping in local optimum, differential evolution (DE) operator has been introduced into PSO algorithm (Wen-Jun and Xiao-Feng, 2003). In this modified PSO algorithm, PSO keeps the particle swarm dynamics at the same time maintains population diversity with the differential evolution.

After the *pbest* and *gbest* of the swarm have been obtained, we will implement a DE mutation operation on the *gbest*. Generate a random r within $[0, 1]$, and a random integer value k within $[1, D]$. When r is less than the given CR or the dimension d is equal to k, and then we can obtain the mutated point *Trail*, whose d th dimension is

$$Trail_{id} = p_{gd} + F \cdot (p_{1d} + p_{2d} - p_{3d} - p_{4d}) (\text{if } r < CR \text{ or } d = k) \quad (3)$$

where CR is a crossover constant, and the scaling factor F is a constant within [0, 2], p_{1d} p_{2d} p_{3d} p_{4d} are chosen randomly from *pbest.*

Then the fitness of Trail can be calculated. Trail will replace p_i only if it is better than p_i. Otherwise, the previous p_i will be retained. This condition is thought to be unsuccessful mutation, that is, new mutation of p_i will resume until the particle skip out of present value.

2.3. Forecasting procedure of improved GPR based on modified PSO algorithm (PSO-GPR)

The forecasting procedure of our method mainly consists of three steps: Step1. Initialization; Step 2. Hyper parameters optimization; Step 3. Prediction. In Step1, we construct the training set and choose the covariance function for Gaussian process regression; and in Step2, the hyper parameters will be optimized based on improved PSO algorithm. The fitness function was set as

$$fitness(\Theta) = -\frac{1}{2} \log \det C_N - \frac{1}{2} T_N^T C_N^{-1} T_N - \frac{N}{2} \log 2\pi \quad (4)$$

in Step3, predictive distribution can be calculated and output the predictions (Fig. 1).

Step 1 (Initialization): Construct the training set and choose the covariance function for Gaussian process regression.

Step 2 (Optimization of Hyper parameters):

Step 2.1 Set parameters for PSO: the size of swarm is *popsize*, the maximum times of iteration is $iter_{\max}$, the accuracy error is ε, $c_1 = c_2 = 2$, $w_{\max} = 0.9$, $CR = 0.9$, $F = 0.9$. Initialize the velocity v_i and position x_i for each particle

Step 2.2 Evaluate fitness of each particle, and find out *pbest* and *gbest*

Step 2.3 Generate a random number r within $[0, 1]$, and a random integer k within $[1, D]$. Implement mutation operation on particles, the *dth* dimension mutated coordinate is $Trail_{id} = p_{gd} + F \cdot (p_{1d} + p_{2d} - p_{3d} - p_{4d})$

Step 2.4 Update the velocity v_i and location x_i of the particles according to equation (2) and (3) respectively. And update the inertial weight by $w = w_{\max} - iter*\frac{w_{\max} - w_{\min}}{iter_{\max}}$.

Step 2.5 Repeat until a stopping criterion is satisfied. That is to say, the process will be concluded when it reach the maximum times of iteration or the improve of the fitness of *gbest* is no more than the accuracy error ε.

Step 3 (Prediction): Input x_* and calculate the predictive distribution of Gaussian is $P(y_* | D, x_*)$. Then output y_*

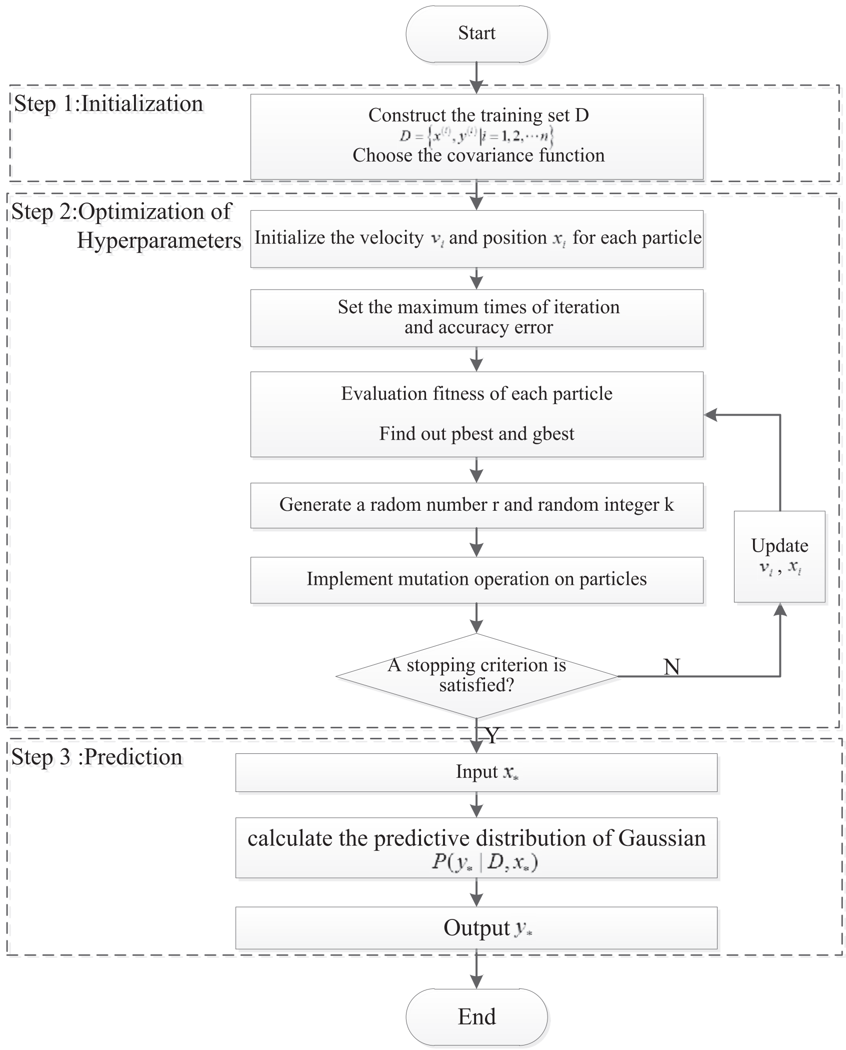

Fig. 1. Flow chart of improved GPR based on modified PSO algorithm (PSO[HYPHEN]GPR).

3. Data description and empirical comparison results

3.1. Data collection

We collected data from U.S. Energy Information Administration; their dataset allows us to find the total carbon dioxide emissions from the consumption of energy for worldwide countries from year 1980–2012, and these data is the latest data as of 2016.

According to the "Trends in Global Carbon Dioxide (CO2) emissions: 2013 Report" (Trends in Global CO2 Emissions: Report 2013, 2013), the six largest emitting countries/regions (with their share in 2012 between brackets) were: China (29%), the United States (15%), the European Union (EU271) (11%), India (6%), the Russian Federation (5%) and Japan (4%). So we selected U.S. and China as our study objects because U.S. and China are the two world's largest economies, and the emissions of the two countries are more than 40% of the world. In addition, Japan, as a representative of the Asian developed countries, early on the implementation of energy-saving emission reduction policies. In 2008, Japan Voluntary Emissions Trading Scheme (J-VETS) was carried out and achieved remarkable emission reduction effect. In 2010, Tokyo implemented the restricted carbon trading system, becoming the first Asia's city which implemented the carbon trading system. So we choose Japan as study object to discuss the Carbon Dioxide emissions trend. Their carbon dioxide emissions from year 1980–2012 are showed in Fig. 2.

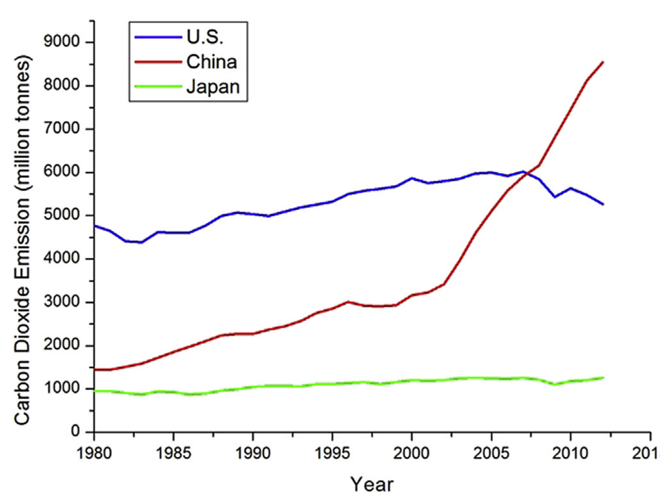

Fig. 2. Carbon dioxide emissions trend of U.S., China and Japan from 1980 to 2012.

Table 1
CO_2 emission predictions in 2003–2012 of U.S. by PSO-GPR, GPR and BPNN.

Year	Real CO2 Emission	Prediction by PSO-GPR			Prediction by3-years lagged GPR	Prediction by BPNN
		1-year lagged	2-years lagged	3-years lagged		
2003	5852.68	5777.0	5786.3	5798.6	5873.706	5948.1
2004	5974.38	5815.7	5823.8	5827.8	5871.109	5816.1
2005	5999.14	5898	5906.1	5903.5	5835.668	5786.2
2006	5923.59	5913.9	5907.9	5898.1	5850.57	5758.3
2007	6024.10	5864.5	5850.5	5845.9	5885.963	5819.5
2008	5840.55	5929.6	5931.6	5922.8	5832.292	5893.7
2009	5429.80	5807.1	5782.9	5781.3	5913.438	5799.8
2010	5630.02	5481.1	5450.3	5469.9	5815.001	6078.3
2011	5483.21	5648.1	5660.2	5691.6	5722.465	6098.7
2012	5270.42	5527.0	5527.1	5514.0	5821.312	5882.2
MAPE		2.74421	2.770449	2.750834	3.548874	5.2535

3.2. *Comparison results of PSO-GPR, original GPR and BP neural networks*

Then we predict the carbon dioxide emissions of U.S., China and Japan in the years from 2003 to 2012 by using the carbon dioxide emissions data from year 1980–2002 as training data. For testing the performance of PSO-GPR, we modeled 1-year lagged, 2-years lagged and 3-years lagged Gaussian process regression models to predict the carbon dioxide emissions of the three countries. Also, we predicted the carbon dioxide emissions by BP Neural Networks with the same data. In addition, we calculated the Mean Absolute Percentage Error (MAPE) for the prediction results, where $MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - x_i}{x_i}\right| \cdot 100\%$, x_i is the real carbon dioxide emission date and y_i is the corresponding predicted value, to compare the performances of 1-year lagged, 2-years lagged and 3-years lagged PSO-GPR, original GPR and BP Neural Networks methods. The prediction results for carbon dioxide emission are showed in Tables 1–3, and the Figs. 3–5 will make the comparisons between the prediction performances of 3-years lagged PSO-GPR, original GPR and BPNN model a little more visually. We only plot the prediction results of 3-years lagged PSO-GPR just because it can represent the average performance of our PSO-GPR in these three experiments.

From Table 1 and Fig. 3, we can find that our PSO-GPR method obtain the best prediction results among PSO-GPR, original GPR and BP Neural Network methods, and the 1-year lagged PSO-GPR model reach the smallest MAPE which is 2.74421 among these forecasting method. It indicates that PSO-GPR method can accurately predict carbon dioxide emissions for the countries like U.S.

Table 2
CO2 emission predictions in 2003–2012 of China by PSO-GPR, GPR and BPNN.

Year	Real CO2 Emission	Prediction by PSO-GPR			Prediction by3-years lagged GPR	Prediction by BPNN
		1-year lagged	2-years lagged	3-years lagged		
2003	3959.97	3507.8	3503.1	3473.9	3496.1	3769.207
2004	4596.97	4033.9	4028.3	3949.2	4012.8	4124.471
2005	5116.35	4647.7	4658.6	4565.1	4617.0	4212.036
2006	5575.20	5140.0	5177.3	5075.2	5031.4	4222.751
2007	5908.43	5568.0	5631.2	5494.0	5300.3	4223.06
2008	6166.57	5874.4	5960.9	5800.7	5436.9	4223.159
2009	6816.10	6109.1	6213.5	6022.5	5491.5	4223.239
2010	7446.52	6688.8	6824	6527.6	5730.1	4223.036
2011	8126.69	7235.5	7415.2	7076.1	5863.8	4223.042
2012	8547.75	7806.2	8033.1	7588.5	5764.9	4223.038
MAPE		9.1322	7.9995	10.7189	16.8946	29.70273

Table 3
CO2 emission predictions in 2003–2012 of Japan by PSO-GPR, GPR and BPNN.

Year	Real CO2 Emission	Prediction by PSO-GPR			Prediction by3-years lagged GPR	Prediction by BPNN
		1-year lagged	2-years lagged	3-years lagged		
2003	1249.72	1204.8	1206	1215.5	1182.626	1154.1
2004	1256.26	1250.3	1249.6	1253.2	1177.979	1149.2
2005	1241.26	1256.3	1257	1248.5	1187.237	1164.2
2006	1239.74	1242.6	1244.2	1248.4	1187.863	1158.5
2007	1254.44	1241.3	1242.4	1252.6	1187.142	1150.7
2008	1216.25	1254.6	1255.1	1260.9	1186.495	1151.6
2009	1104.91	1219.8	1222.4	1225.5	1187.926	1162.9
2010	1180.58	1116.7	1122.8	1144.8	1119.584	1161.6
2011	1200.27	1187	1185.3	1223.5	1156.207	1053.3
2012	1259.06	1205.1	1205.8	1190.8	1178.829	1163.6
MAPE		3.0914	3.0818	2.9380	5.0682	6.9203

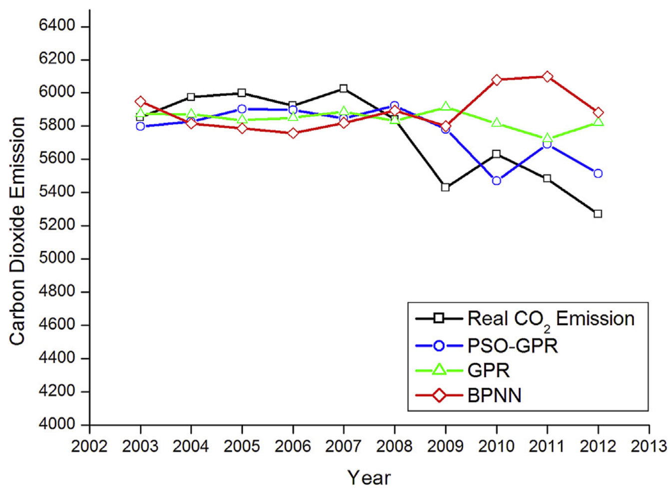

Fig. 3. Carbon dioxide emission predictions of U.S. by PSO-GPR, GPR and BPNN.

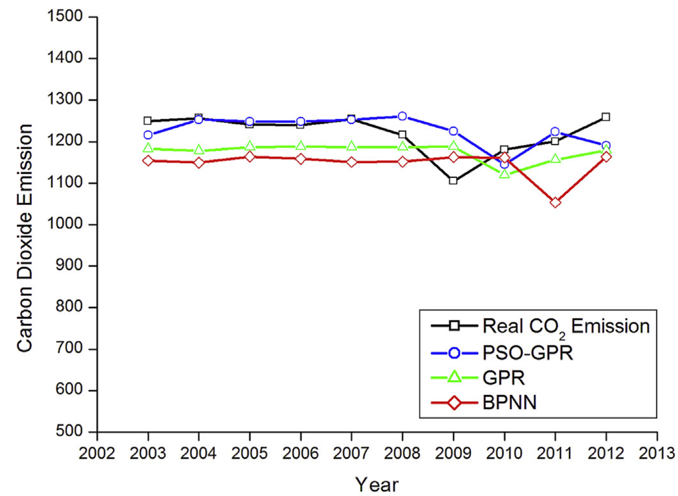

Fig. 5. Carbon dioxide emission predictions of Japan by PSO-GPR, GPR and BPNN.

From Table 2 and Fig. 4, it can be observed that 2-years lagged PSO-GPR works best and the accuracy of BP Neural Networks became very worse, although all the MAPE of PSO-GPR and BP Neural Networks are much larger than the MAPE in experiments of U.S. Maybe the reason is that the carbon dioxide emission of China rise dramatically after 2000, but we use the carbon dioxide emission data from year 1980–2002 as training data in our machine learning methods, hence the parameters we learning from the training data can't be on behalf of the trend of China in 2003–2012. So we had better use a resent and short timing series data as training data to predict the carbon dioxide emission of the kind of countries just like China whose carbon dioxide emission increase very rapidly.

Also, PSO-GPR performed better than BP Neural Networks in predicting Japan Carbon Dioxide emission. The MAPE of 3-years lagged PSO-GPR is smallest. The Fig. 5 shown that the gap between real carbon dioxide emission and PSO-GPR predictions became larger after 2007 and became narrow after 2010. The reason is mostly due to the Kyoto Protocol which limit Japan's Carbon emission during this certain period and the carbon dioxide emissions of Japan in these years have been reduced largely.

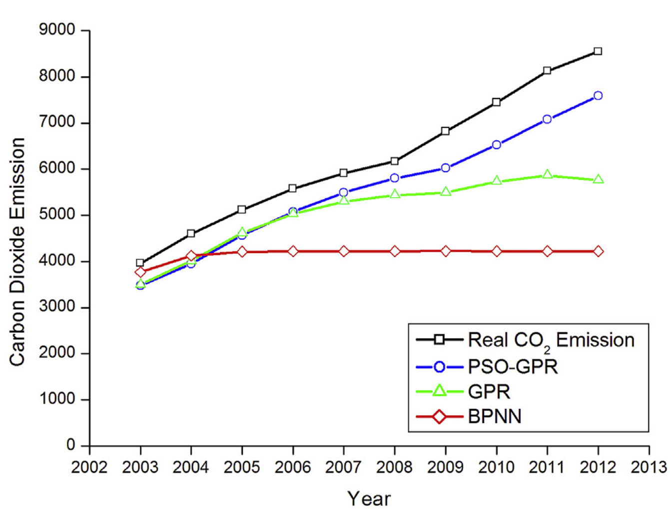

Fig. 4. Carbon dioxide emission predictions of China by PSO-GPR, GPR and BPNN.

3.3. Application PSO-GPR to predict total carbon dioxide emissions by 2020

We applied PSO-GPR method in generating a prediction total carbon dioxide emissions for 2013 to 2020 as Fig. 6 shows. The predictions are based on the total carbon dioxide emissions data of U.S., China and Japan from 1980 to 2012 from U.S. Energy Information Administration.

Our prediction demonstrates that after 2012 China's total carbon dioxide emission will continue to increase at an increasing rate for three to four and then increase at a decreasing rate after 2015, and the increase at a decreasing rate is partly due to the Paris agreement accords (Lewis, 2016). In December 2015, China joined United Nations Framework Convention on Climate Change (UNFCCC), which required all parties to promote five key elements: long-term goals, climate funding, action, transparency, and adaptation. And this agreement effectively promoted China's carbon emissions rate decreased (Elzen et al., 2016). Japan has a tiny downward line after 2012 and the United States curve seems convergence around 5650 metric tons. The results from PSO method support the idea that China's total carbon dioxide emission will still increasing but finally will be at a decreasing rate, and the developed countries, such as U.S. and Japan will have a good control on their amount of carbon emission in the near future.

Under the background of the new normal, China's carbon dioxide emissions continued to grow until reach the inflection point, then began to decline, and in 2020 is going to reach about 103.34 tons of total carbon dioxide emissions. Compared with the developed countries such as the U.S. and Japan, the growth rate of China's carbon dioxide emissions is faster. The main reason is the rapid growth of China's economic development and the transformation of the process of urbanization. All of these promotes the increase of carbon dioxide emissions of China.

China is still a developing country, and economic development is the primary problem at present. It is generally believed that the way of reducing total carbon dioxide emissions is to control industrial investment and energy consumption. China, as an industrializing nation, carbon dioxide emission reduction will prevent economic growth partly, so the government should deal with the relationship between economic growth and carbon dioxide reduction, and economic development should not be overwhelmingly influenced by carbon dioxide emission reduction policy.

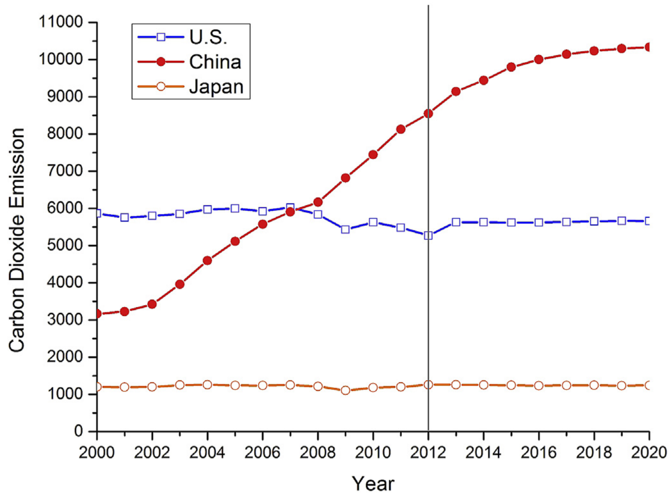

Fig. 6. Carbon dioxide emission predictions of China, U.S. and Japan by PSO-GPR for 2013 to 2020.

In addition, there is a great gap between the regions' economic and environment in China. Therefore, in the process of urbanization, the government should implement different carbon dioxide emissions reduction policy in different regions. For example, some of the carbon dioxide emissions heavily regions, such as the northeast China and the Beijing-Tianjin-Tangshan (BTT) regions, should be regarded as the key area of implementing carbon dioxide emission reduction work. In contrast, for these regions whose economic development is relatively slow, and the local environment has a remaining capacity, the government should support its economic development, and should not put too much pressure to reduce carbon dioxide emissions.

4. Conclusions

In this paper, we endeavor to make the forecasting method more efficient and effective for CO_2 emission prediction. Our major contribution is improving the Gaussian process regression method by optimizing the hyper parameters in the Gaussian process regression model. Also we tested our improved GPR method with the total carbon dioxide emissions data of U.S., China and Japan in 1980–2012. And then, we have compared the prediction performances of our method with BP Neural Networks by the data of U.S., China and Japan.

The U.S., China and Japan's cases support the idea that PSO-GPR is a better method in carbon dioxide emission prediction. The results show that the MAPE of predictions generated by PSO-GPR are less than that from original GPR and BPNN method. That means our PSO-GPR is an enhanced method for CO_2 emission forecasting and the performance of PSO-GPR is superior to original GPR method and the traditional method like BP Neural Networks. In addition, from comparing of the MAPE of 1-year lagged, 2-years lagged and 3-years lagged Gaussian process regression models, we found that the performance of the models is depend on whether the variation tendency of CO_2 emissions is stable. If the CO_2 emissions fluctuated wildly or changed rapidly, 1-year lagged PSO-GPR will perform better, otherwise 3-years lagged PSO-GPR will be more suitable.

For further research venues, we will expand our model to apply it to more other countries, especially developing countries because these countries are more vulnerable to climate change and might sacrifice sustainable environment development in exchange for economic growth. We will take into account other effect factors in our PSO-GPR method, such as GDP and populations as well. Furthermore, the estimation results show that Kyoto Protocol has played an important role in the carbon dioxide emission reduction and inspired us that environment policy also need to be considered in the carbon dioxide emission forecasting method in future research.

Conflicts of interest

The authors declare no conflict of interest.

Kennedy and Eberhart, 1995; Liu et al., 2010; Morita et al., 1994; Tobias et al., 2006; Zhang et al., 2013.

Acknowledgments

This research was financially supported by the National Natural Science Foundation of China (Grant No. 71673210, 71373198, 71231007, 91647119). The authors would like to thank the funded project for providing material for this research. We would also like to thank our anonymous reviewer for the valuable comments in developing this article.

References

Abdel, Karim Baareh, 2013. Solving the Carbon Dioxide emission estimation problem: an artificial neural network model. J. Softw. Eng. Appl. 6 (7), 338–342.
Andrew, J.W., Gillian, K., Jemma, L.G., Richard, A.B., 2010. The impact of climate, CO2 and population on regional food and water resources in the 2050s. Sustainability 5 (5), 2129–2151.
Piero, Baraldi, Francesca, Mangili, Enrico, Zio, 2015. A belief function theory based approach to combining different representation of uncertainty in prognostics. Inf. Sci. 303, 134–149.
Martin, Bogdan, Dominik, Brugger, Wolfgang, Rosenstiel, Bernd, Speiser, 2014. Estimation of diffusion coefficients from voltammetric signals by support vector and gaussian process regression. J. Cheminformatics 6, 30.
Chen, Junghui, Teck Chan, Lester Lik, Cheng, Yi-Cheng, 2013. Gaussian process regression based optimal design of combustion systems using flame images. Appl. Energy 111, 153–160.
Dell, Melissa, Benjamin, F., Jones, Benjamin, A., Olken, Y., 2013. Temperature shocks and economic growth: evidence from the last half century. Am. Econ. J. Macroecon. 4 (3), 66–95.
Deschenes, Olivier, Greenstone, Michael, 2007. The economic impacts of climate change: evidence from agricultural output and random fluctuations in weather. Am. Econ. Rev. 97 (1), 354–385.
Dragomir, V.D., 2012. The disclosure of industrial greenhouse gas emissions: a critical assessment of corporate sustainability reports. J. Clean. Prod. 30 (5), 222–237.
Elzen, M.D., Admiraal, A., Roelfsema, M., et al., 2016. Contribution of the G20 economies to the global impact of the Paris agreement climate proposals. Clim. Change 137, 1–11.
Feng, Shuaizhang, Krueger Alan, B., Michael, Oppenheimer, 2010. Linkages among climate change, crop yields and Mexico–US cross-border migration. Proc. Natl. Acad. Sci. 107 (32), 14257–14262.
Geng, Y., Zhao, H., Liu, Z., Xue, B., Fujita, T., Xi, F., 2013. Exploring driving factors of energy-related CO2 emissions in Chinese provinces: a case of Liaoning. Energy Policy 60, 820–826.
Mikiko, Kainuma, Yuzuru, Matsuoka, Tsuneyuki, Morita, 2000. The AIM/end-use model and its application to forecast Japanese carbon dioxide emissions. Eur. J. Operational Res. 122, 416–425.
Kavoosi, H., Saidi, M.H., Kavoosi, M., Bohrng, M., 2012. Forecast global Carbon Dioxide emission by use of genetic algorithm. Int. J. Comput. Sci. 9 (5), 418–427.
Kennedy, J., Eberhart, R., 1995. Particle swarm optimization. In: IEEE Int. Conf. On Neural Networks, vol. 4, pp. 1942–1948.
Kou, Peng, Gao, Feng, Guan, Xiaohong, 2013. Sparse online warped Gaussian process for wind power probabilistic forecasting. Appl. Energy 108, 410–428.
Kyle, W.K., Juliet, B.S., 2014. Economic growth and climate change: a cross-national analysis of territorial and consumption-based carbon emissions in high-income countries. Sustainability 6 (6), 3722–3731.
Lewis, S.L., 2016. The Paris Agreement has solved a troubling problem. Nature 532, 283.
Liu, Lan-Cui, Wang, Jin-Nan, Wu, Gang, Wei, Yi-Ming, 2010. China's regional carbon emissions change over 1997-2007. Energy Environ. 1 (1), 161–176.
Mattinen, M.K., Heljo, J., Vihola, J., et al., 2014. Modeling and visualization of residential sector energy consumption and greenhouse gas emissions. J. Clean. Prod. 81 (7), 70–80.
Robert, Mendelsohn, Nordhaus, William D., Shaw, Daigee, 1994. The impact of global warming on agriculture: a Ricardian analysis. Am. Econ. Rev. 84 (4), 753–771.
Miao, Z., Geng, Y., Sheng, J., 2016. Efficient allocation of CO2 emissions in China: a

zero sum gains data envelopment model. J. Clean. Prod. 112, 4144–4150.
Morita, T., Matsuoka, M., Penna, I., Kainuma, 1994. Global Carbon Dioxide Emission Scenarios and Their Basic Assumptions: 1994 Survey. Center for Global Environmental Research, Tsukuba, Japan. CGER-I011-'94.
Remya, Nair, Sanjay, Jhingan, Deepak, Jain, 2015. Testing the consistency between cosmological measurements of distance and age. Phys. Lett. B 745, 64–68.
Nordhaus, William, D., 2006. The "Stern Review" on the Economics of Climate Change. National Bureau of Economic Research. No. w12741.
Arash, Pourhabib, Faming, Liang, Yu, Ding, 2014. Bayesian site selection for fast Gaussian process regression. IIE Trans. 46, 543–555.
Rasmussen, C.E., Williams, C.K.I., 2006. Gaussian Processes for Machine Learning. MIT Press.
Reijnders, L., Huijbregts, M.A.J., 2008. Palm oil and the emission of carbon-based greenhouse gases. J. Clean. Prod. 16 (4), 477–482.
Stern, Nicholas, 2008. The economics of climate change. Am. Econ. Rev. 98 (2), 1–37.
Sun, Wei, 2013. Forecasting Carbon Dioxide emissions in China using optimization grey model. J. Comput. 8 (1), 97–101.
Sun, W., Liu, M., 2016. Prediction and analysis of the three major industries and residential consumption CO2 emissions based on least squares support vector machine in China. J. Clean. Prod. 122, 144–153.
Tobias, P., Malte, K. and Carl, E. R, 2006. Nonstationary gaussian process regression using a latent extension of the input space, ISBA Eighth World Meeting on Bayesian Statistics.
Trends in Global CO2 Emissions: Report 2013, 2013. PBL Netherlands Environmental Assessment Agency.
Wang, Jianzhou, Hu, Jianming, 2015. A robust combination approach for short-term wind speed forecasting and analysis–-combination of the ARIMA (Autoregressive integrated Moving average), ELM (extreme learning machine), SVM (support vector machine) and LSSVM (least square SVM) forecasts using a GPR (gaussian process regression) model. Energy 93, 41–56.
Wen-Jun, Z., Xiao-Feng, X.D.E.P.S.O., 2003. Hybrid particle swarm with differential evolution operator. In: IEEE International Conference on Systems, Man & Cybernetics (SMCC), vol. 4, pp. 3816–3821.
Yu, Jie, Chen, Kuilin, Rashid, Mudassir M., 2013a. A Bayesian model averaging based multi-kernel Gaussian process regression framework for nonlinear state estimation and quality prediction of multiphase batch processes with transient dynamics and uncertainty. Chem. Eng. Sci. 93, 96–109.
Yu, Jie, Chen, Kuilin, Mori, Junichi, Rashid, Mudassir M., 2013b. A Gaussian mixture copula model based localized Gaussian process regression approach for long-term wind speed prediction. Energy 61, 673–686.
Zhang, Xiaoling, Shen, Geoffrey Q.P., Feng, Jingjun, et al., 2013. Delivering a low carbon community in China: technology vs. Strategy? Habitat Int. 37 (1), 130–137.